Actions Speak Louder Than Words: The Production and Enactment Effects on Recall and Recognition Memory of Verbs

Hayley Blowfield

A report submitted as a partial requirement for the degree of Bachelor of Psychological Science with Honours in Psychology at the University of Tasmania, 2018

University of Tasmania

2018

Statement of Sources

I declare that this report is my own original work and that contributions of others have been duly acknowledged.

______________________________ Date:______________________

Hayley Blowfield

Acknowledgments

I am pleased to acknowledge the many people who have helped me throughout the process of completing this challenging and gratifying research. Firstly, I would like to thank all the people who took the time to participate in my study. If it were not for all of them, I would have had no data to write about! I really appreciated the effort many of them took to make it to the experiments, even under such short notice.

A huge thanks my supervisor, Dr Mike Garry, for his guidance, and giving so much of this year to assist with all aspect of me completing my thesis research. His patience in answering my nonstop questions, his attention to detail in providing hard to hear feedback, and his sensible wit in our thesis meetings was greatly appreciated throughout this entire experience. I would also like to thank my honours colleague's, Zac Bain-Williams and Michael O'Leary. Bouncing ideas off my thesis team, and their additional input, often helped me critically think about my own research and ways to improve upon it.

Nick and Jobelle, two of the closest people in my life, I appreciate you both so much. If it were not for your support, care, and unwavering belief in me this year, I am not sure I would have completed this research this year. I owe the completion of this thesis to you both. Jobelle, you helped keep me focused and motivated over the year. Nick, I think you have re-read my thesis over 50 times by now. Thank you for checking over me and my work time and time again.

I would finally like to thank my Mum, and two sisters, Jess and Sarah. They have been such a help, especially in the final weeks of writing this thesis. I cannot explain how much I appreciate the effort that they put into giving me guidance and support, despite being so far away from Tasmania, and none of them having ever come close to reading or writing a research report.

Table of Contents

List of Tables and Figure

Actions Speak Louder Than Words: The Production and Enactment Effects on Recall and Recognition Memory of Verbs

Hayley Blowfield

Word Count: 9588

Abstract

This study aimed to investigate the effects of gestural and verbal production of single verbs on memory and expected to find that the physical embodiment of verbs via miming would significantly increase memory retention. It was also expected that gesturing verbs while speaking them aloud would lead to significantly more produced verbs being remembered than when they were only gestured. 12 participants (6 female; 1 left-handed) aged 18-27 were recruited from the University of Tasmania and via snowball sampling and were randomly assigned to either the gesture only group, or the gesture/spoken group, where they studied 80 verbs coloured blue or white which indicated whether to produce or silently read the words. Contrary to what was expected, there were no significant between subject effects, suggesting no benefit to vocalising verbs whilst performing them as gestures. However, there were extremely large within subject effects, suggesting that mimed performance of verbs significantly increases memory. These results suggest that the enactment of verbs leads to significant memory benefit, while the addition of speaking aloud verbs may not lead to any additional effect. However, further investigation is necessary to address the study's limitations.

Keywords: Gesturing, Reading, Learning, Free Recall Memory, Recognition Memory.

Slamecka and Graf (1978) discovered that generating a word from a cue increased word learning, compared to simply reading a presented word silently. They presented a word stimulus (e.g. rapid) to participants, and then asked them to generate a related word response (e.g. synonym), beginning with a given letter (e.g. 'F'). Participants responded by generating 'fast', as the correct synonym to 'rapid' beginning with 'f'. These generated words were better remembered than words read. This was adeptly named the generation effect and led to the discovery that information is better remembered when generated by one's own mind, rather than simply presented.

Hopkins and Edwards (1972) were the first to test whether pronouncing words led to better recognition than reading words and found this to be case in a within subject, mixed-list designs. That is, participants would remember more spoken words from list of word they studied by both speaking and silently reading words. However, this finding had not been researched again until recently. MacLeod et al. (2010) further investigated this production effect and replicated the initial results, finding that a word is learned more effectively if spoken aloud when compared to words read silently. An effectively large and consistent memory advantage has since been shown for reading aloud over reading silent, which has shown across a wide variety of conditions. Reading aloud, or verbally producing, has a greater effect than hearing words being spoken (MacLeod, 2011), mouthing words (MacLeod, et al., 2010), writing words (Dewhurst, Rackie, & Esch, 2016; Forrin, MacLeod, & Ozubko, 2012), speaking words quietly (Quinlan, & Taylor, 2013), and typing words (Forrin, et al., 2012) when they were studied alongside words that were read silently.

The enactment effect (Engelkamp, Zimmer, & Kurbjuweit, 1995), or self-performed task effect, has also been established in the literature as an effect where reading a brief instruction and then responding by performing a related action leads to better memory of the instruction than just listening to, or reading it. Participants who embody an instruction (e.g. close the door) are more likely to remember the instruction than if had they simply listened to them being spoken.

All three of these encoding enhancing effects involve overt responses unique to each stimulus, found in a within-subject mixed-list paradigm, which ultimately leads to enhanced responses to explicit recognition memory tests. Researchers have since been expanding on these effects across a range of modes of presentation, production, and recall methods.

Theoretical Underpinnings of Memory Effects

There have been two predominant theories that have endeavoured to explain the underlying mechanics of these memory effects. The distinctiveness theory suggests that information attended to is remembered by recalling distinctive aspects of the item and tracing the information back to the source. The more distinctive an item is, the more likely it will be remembered over less distinctive items. The strength account instead suggests that information that is attended to will be more likely recalled if the item was strongly attended to, elaborated on, and consolidated. The strength account supports the original production effect hypothesis that people who speak or perform words would remember more words than people who only read them silently. However, this has not been predominantly found to be the case with the production effect. So, the distinctiveness account helped explain why the production effect has mainly only been observed in within-subject, mixed-list designs. This is where the spoken or performed words were remembered only when there were words within the list that were read silently, as this makes the produced words more distinctive than the baseline silent words.

The distinctiveness account. The production effect has been found to create greater benefit to recognition memory for produced words than unproduced words. This has been largely attributed to the distinctiveness account (MacLeod, et al., 2010). This suggests that the production effect can be accounted for at encoding; that verbal production creates additional mnemonic information stored as a memory trace that can be reactivated, or tapped into, at test to discriminate produced words from unproduced words. People may remember saying an item aloud and infer that is was studied. As distinction is relative, this trace can only be created when the spoken words are distinct to a baseline of undistinctive words (silently read words), which explains why large effects are only found in mixed-list designs (MacLeod, et al., 2010). This also suggests that the memory retrieval strategies used at test are linked to the distinctive production process at study (Fawcett, 2013).

The strength account. A competing explanation to the distinctiveness account is the strength account; where the strength of the studied item's representation in memory alone predicts recognition (Ozubko, & MacLeod, 2010). The strength account suggests that producing an item creates a stronger, more solid and condensed memory, making it easier to retrieve and recognise than weaker items that are read silently. This may seem to account for the 'lazy reading hypothesis' (Forrin, Jonker, & MacLeod, 2014), suggesting that words read aloud often compromise the memory retention of silently read words, and that the memory benefit to the spoken words are often to the detriment of silent words. This detriment is said to be due to lack of elaborative encoding of silently read words, rather than not having a unique feature to encode and later remember, as the distinctiveness account suggests. However, MacLeod et al. (2010), found evidence against this hypothesis, when they enforced semantic processing for both read aloud and silently read words, and still found a production effect (also see, Forrin, Jonker, & MacLeod, 2014).

Likely a combination of both theories. The production and enactment effects are therefore likely due to a combination of both strength and distinctiveness that are not mutually exclusive. While many individual studies of the production effect have not found any group differences in rates of recognising words, despite one group saying all words aloud rather than silently, a meta-analysis revealed a consistently small between-subjects effect that supports the strength account (Fawcett, 2013). There has also been a large within subject effect consistently found, suggesting that words produced are more distinctive and therefore easier to recall when they have been studied with silently read words. As a small between subject production effect has been established, this may suggest that the relevant memory traces created by distinctiveness (via mixed-list study) are also strengthened by elaborative production, and therefore improves recognition as a single process (Fawcett, 2013).

The distinctiveness account only suggests it is the additional memory trace that is used at test to recall and disregards the role of encoding strength. This would mean two separate groups, which each were studying a mixed-list, but each group were using a different mode of production (vocal vs. typing) would be equally distinctive and therefore would create equal memory traces. If they had equal memory traces, as the distinctiveness account suggests, this would mean they would each lead to the same production size effect. However, this has not been found to be the case, and vocal production is more effective than typing (MacLeod, et al., 2010; Forrin, et al., 2012).

Therefore, it is suggested that ‘strength’ plays an important role in within-subject design by strengthening these distinctive traces created by producing in a unique way, improving recognition on a mode of production basis. For example, speaking aloud will produce a stronger distinctiveness trace than mouthing or whispering, because speaking requires more cognitive effort, which increases the strength of the distinctive traces created by producing. The strength of the semantic encoding leads to a small production effect and is coupled with the distinctive memory traces and further strength of encoding due to more attracted attention to words spoken aloud which produces the well-established large within subject production effect (Fawcett, 2013).

Principle theory of embodied cognition. The theory of embodied cognition is an overarching idea and foundation throughout cognition research. Embodied cognition suggests that cognition is shaped by the body of the organism experiencing it. For humans, we can experience quite complex cognition, which is impacted by our multifaceted motor functioning, the sensitivity of our senses and perceptual systems, and our bodily interaction with the world around us. “Sensorimotor information is simulated during language processing” (Sidhu, & Pexman, 2016), and differences in this simulation can affect the success of word learning and memory. For example, verbs that have more specific bodily information and therefore more embodied meaning (e.g. ‘lunge’ compared to ‘work’) facilitate lexical processing (Sidhu, & Pexman, 2016).

This has not been theorised much since by previous researchers of the production effect, which has a focus of verbal production of nouns. However, the enactment effect which involves the body more readily through the bodily enactment of actions phrases, leads to a 20-30% increase in recognition rates (Engelkamp, et al., 1995). This effect is greater than that of the production effect, which only increase recognition by 10-20% (Engelkamp, et al., 1995). This is likely due to the direct embodiment of the verb action phrases and their meaning, which has a direct impact on the effectiveness and strength of the cognition effect. As verbs have more embodied information, processing them would facilitate simulation of sensorimotor information (Sidhu, Kwan, Pexman, & Siakaluk, 2014). However, Sidhu, and Pexman (2016) did find that embodiment and enactment effects do not interact, which suggests that the memory benefit of embodiment simulation, and memory benefit of deliberate imagined enactment simulation, may be separate unrelated processes.

Will verbs create larger enactment and production effects than nouns? As all previous production effect research looks at noun retention, a production effect of verbs may produce larger memory benefits in memory than the current production effect size, which has been consistently found to be around 10-20% (MacLeod, et al., 2010). Assuming embodied cognition is indeed related to this effect, this increase in production effect size would be due to verbs being more related to movement and body, than nouns, which are more visually involved via mental imagery. Furthermore, if verbs do increase memory retention, this may also account for some of the increased sizes of effect of previous verb enactment effect studies, compared to noun production effect sizes.

Review of Previous Research

The production effect persists alongside the generation effect. A study by MacLeod (2011) found that saying words aloud led to better recognition than listening to the words being read aloud by someone else. This gives evidence to the generation effect, as it found that a person who generates a word themselves via production, will lead to better memory than if simply presented to them. It also suggests that the self-referencing effect also plays an important role in the magnitude of memory effects (also see, Forrin, & MacLeod, 2017).

MacLeod, et al. (2010) found that melding the generation effect and production effect together led to greater overall accuracy of recognition, as well as higher overall percent word recognition. This showed that adding verbal production to other known memory effects can establish greater distinctiveness than the original memory effect alone. Additionally, this gives further evidence to the predominant underlying theory, the distinctiveness account, by suggesting that production is as beneficial to words strongly encoded (with the addition to the generation effect) as to words weakly encoded (traditional production effect paradigm).

Benefits of using the production effect. Researchers have determined that the production effect is robust and has a consistent benefit to memory of nouns over time. Ozubko, Hourihan, & MacLeod (2012) found that the production effect led to memory benefit even after one week without overt elaboration or restudying. It has also been found that continuous enactment can lead to substantial and practical memory benefits (Steffens, Stulpnagel, & Schult, 2015). Many mnemonic devices exist in the literature, such as elaboration (Craik, & Lockhart, 1972), translation (Conway, & Gathercole, 1990), drawing (Wammes, Meade, & Fernandes, 2016) and generation (Slamencka and Graf, 1978). While these effects sizes may rival that of the production effect (10-20%) and enactment effect (20-30%), none are as easy to implement and execute in learning strategies as producing and enacting. (Ozubko, Major, & MacLeod, 2014).

When does the production effect work, and not work? MacLeod, et al. (2010), in their endeavours to expand on the production effect literature, found that saying yes or pressing the 'space' bar on a keyboard when presented with items as a mode of production, did not increase memory retention when compared to reading an item silently. This suggests that it is not solely the overt act of moving (by pressing) nor the vocal production that leads to the production effect. Rather, it is the additional unique and meaningful encoding features that create a distinctiveness in items that ultimately leads to the production effect.

To further demonstrate this point, MacLeod, et al. (2010) found that mouthing words led to a significant, but smaller than vocalising, production effect, again suggesting that vocal output alone is not what creates distinctive differences to silently read words. However, it was found that non-words, which lack any semantic meaning, led to a large production effect when verbally produced, despite overall lower recognition rates due to unfamiliarity (MacLeod, et al., 2010). This could be due to the vocal pronunciation being unique to each non-word produced, which created distinctive meaningful memories, similar to how the unique semantic meaning of words can also create unique distinction when mouthed.

Utilising elaborative encoding. An example of deeper elaborative encoding of words would involve focused visual imagining, drawing (Wammes, Meade, & Fernandes, 2017), being in contact with an object, or embodying a words meaning as a gesture or movement. Fawcett, Quinlan, and Taylor (2012) found that nouns are better remembered when presented as pictures rather than as words. It has also been found that verbally producing names does not improve face-name association memory (Hourihan, & Smith, 2016). Production seems to only enhance associative memories when both members of pairs are produced. Also, the production effect may only effect associative memories if the items studied are already embedded in long-term memory, unlike the study of new names-faces. This suggests that the more literal the presentation or production of the word, and the more previously embedded studied items are in memory, the more likely items will be remembered at test.

The addition of the production effect. Engelkamp, et al. (1995), found that verb phrase enactment during study led to better memory recognition than listening to the phrases being pronounced, naming this the enactment effect. The enactment effect is quite large by itself (20-30%) and may be due to it involving a more encompassed interaction with encoding material. Coupling the enactment effect with other modes of more simple production, such as speaking aloud, could further increase this memory effect. Verbal production is easily implementable, especially alongside the enactment effect.

The enactment effect on recall is more distinct for phrases when object cues that appear in the phrase are also within the study context, such as a table for studying the phrase, 'knock the table' (Steffens, Buchner, Wender, & Decker, 2007). This is likely due to additional distinctive outer cues that can directly be tapped into during test. Participants can create a distinctive trace for the unique enactment, and use the unique visual object associated with the enactment, as an additional cue to activate the trace. When imagined-object phrases, and interactive-object phrase were separated, it was found that the imagined-object phrase recall was no better than reading a word aloud, and the enactment effect only persisted when interactive objects were present (Steffen, et al., 2007). So, actions acted on objects do speak louder than words. However, it is yet to be determined whether actions that are instead acted on imagined objects, with verbs read aloud would speak loudest or hinder each effect. The former is expected, as many learning experiences such as language acquisition often depends on simultaneous spoken and gestured production.

How experimental design impacts results of these effects. Previous studies have managed to replicate and reproduce the overall finding of a large production effect in within subject mixed-list designs. That is, studies where all participants learn a list of words, using 2 different methods (e.g. verbally producing, and silently reading), will often lead to participants remembering more of the words they produced than the word they read. However, until recently, there was no research suggesting a between-subject pure-list production effect. This refers to studies where one group of people learn a list of words using one method (verbally producing), and another group of people learn the same list with a different method (e.g. silently reading), expecting that the group who produced their list would remember more words than the group who read their list of words.

Castel, Rhodes, & Friedman (2013), like many other small sample studies, found no difference in the amount of words remembered in a between groups production effect study. However, they did find that production significantly improves judgements of learning: a person’s accuracy in judging whether they have successfully learnt an item or not. The enactment effect was found with substantial between subject effects when contrasted to the production effect’s small between subject effects (Engelkamp, et al., 1995). The typical study of the enactment effect uses verb-object phrases (e.g. point to the window), which contextualise the action being performed. The phases request the participant to either use an object to complete the action or imagine the object to complete the task.

But wait! The production effect exists between subjects. Fawcett (2013) conducted the first meta-analysis on the production effect that found a consistent small between subject effects of pure-list studying. This effect had only been reported once before by Gathercole and Conway (1988). The most recent meta-analysis and experiment reported similar findings to Fawcett's (2013) meta-analysis, also revealing a significant but small between-subjects effect (Bodner, Taikh, & Fawcett, 2014). This does not necessarily rule out a distinctiveness account for the between-subjects production effect, as a people who study a list purely with spoken production can actively choose to attempt to inform their recognition decisions, by trying to recollect whether they spoke an item aloud at study. However, it does give stronger evidence for the strength account. There was no difference between the within group spoken word accuracy, and spoken only group word accuracy, as the distinctiveness theory predicts (MacLeod, et al., 2010).

Rationale

Recall and recognition memory. Like the production effect, the enactment effect has previously found inconsistent results on their effects on free recall (Steffens, et al., 2007). Recall memory relies on explicit retrieval of highly valuable information about the item to make an inference, and/or the item itself. However, recognition relies on additional information being provided to activate an otherwise dormant memory trace which can either lead to complete recognition or can lead to a feeling of familiarity, which strength determines a guessed answer to an often-binary question (yes/no) in recognition memory tests. The problem with recognition-only designs is that it can be difficult to determine what levels of the word is recognised; whether they could have been explicitly recalled independent of the cue, complete recollection, or inferred/ partial guesses via familiarity.

Recollection, or just familiarity? One way to measure this is by asking participants to self-report their feelings of familiarity or recollection during recognition tasks. A production effect study conducted by Fawcett, and Ozubko (2016), measured participants source memory by asking them to self-report whether they ‘remember’ or ‘know’ the word; ‘know’ indicating that they remember the specific context the word was presented or produced as. Taikh, and Bodner (2016) have also suggest using self-reports as another means to assess when participants use a distinctiveness strategy to remember an item, such as recognising a word due to its distinctive attributes in the previous task (e.g. produced).

Ozubko, Gopie, and Macleod (2012), found that people more accurately recollect spoken words as being spoken, than they recollect silently read words as being read (also see, Ozubko, Major, & MacLeod, 2014). It was also found that production, in a mixed-list study design, enhances recognition through both recollection and familiarity (Ozubko, Gopie, & MacLeod, 2012). However, it seems recognition of pure-list words, studied by vocal production, is only enhanced through familiarity (Fawcett & Ozubko, 2016). Jones, and Pyc (2013), also found that free recall of produced words is only greater than silent words in a mixed-list paradigm, and this ‘improvement’, was rather a decrease in free-recall-ability for silently read words. This is otherwise known as the list-strength effect, which refers to the phenomena where recall memory benefits of producing a mixed-list is to the detriment of read word recall rates (Ratcliff, et al., 1990; Verde, 2009). Free recallability, familiarity and recollection of source-information is therefore all likely impacted to some extent by both the production and enactment effect.

Aims and Hypotheses

The aim of this study is to investigate the combined effects of enactment and production. As there have been mixed findings on whether verbal production improves recall, as well as recognition, I aim to investigate both systems of memory

(Fawcett, & Ozubko, 2016). Secondary to this, I will also be investigating source memory information to determine whether groups differ in their levels of ‘explicitly known’ or ‘vaguely familiar’ word recognition. I aim to examine single verbs rather than object dependant phrases, and I will be examining recognition accuracy rates, to determine sensitivity in responses, considering correct identification and false alarm recognition rates, to see whether they differ between production groups or word type (produced/ silent).

The proposed research hopes to find that the physical production of verbs via miming will increase memory retention. It is hypothesised that verbs produced with gestures, compared with verbs read silently, will lead to significantly higher recall and recognition. It is also hypothesised that verb recall, and recognition will be greater when verbs are produced simultaneously as gestures and being spoken aloud, contrasted with silently read. Finally, it is hypothesised that participants who simultaneously gesture and speak verbs aloud will have significantly greater recall and recognition of verbs than participants who only gesture verbs.

Method

Participants

A convenience sample of 12 participants (6 female; 1 left-handed) aged 19-27 (M= 22.8, SD= 2.4), were recruited primarily from the University of Tasmania Sandy Bay campus, via posters and online advertisement, as well as through my networks of friends and colleagues. Ethics approval was obtained quite late and meant there was only a small window in which to collect data (Appendix C). Due to this time delay, there was a deficit in participant recruitment. All had met criteria of having corrected/normal vision, no known neurological, neuromuscular, learning, or reading disabilities, having completed at least grade 10, aged 18-50 years, and knew

English as first language. Psychology students received course credit, and all other participants were placed in a draw to receive one of six $100 gift cards.

Materials and Apparatus

120 verbs were selected, 3-8 letters long in the English language, which could be performed using the upper body and face. 80 verbs were used in the experimental study/learning task (Appendix A), and 40 were used as filler verbs in the recognition task (Appendix B). The computer program PsychoPy (Peirce, 2009) was used to present the verb stimuli, and to collect recognition and remember-know judgment data. During the study phase, verbs were presented either in white or blue letters in 5cm high font. The font colour indicated to the participant whether to silently read, or to gesture and/or speak the verb. During the recognition and remember-know judgments task, all verbs were presented in yellow in 5cm high font, to eliminate context clues.

Procedure

After participants were screened, they were invited to attend a single 90-minute session to participate in the study. I obtained informed consent, after the participants had read a study information sheet, and I had explained the study procedures. They were then assigned a unique four-character ID (two random letter and two random digits, e.g., AB12), which was then recorded on the signed consent form.

Verb study phase. Participants sat at a desk in front of a laptop computer. I randomly assigned participants to either the Gesture Group, or the Gesture-Spoken Group and told them that for the first phase of the experiment, verbs would appear on the screen one at a time. I then instructed the Gesture group participants to perform the verbs presented in blue letters as gestures, and to silently read the verbs presented in white letters. I instructed the Gesture-Spoken group to both perform (gesture) and simultaneously speak aloud verbs presented in blue letters, and to silently read verbs presented in white letters. I also requested that all participants not move their lips while silently reading the white verbs, or while gesturing the blue verb if they were part of the Gesture group, as mouthing has been shown to impact memory of verbs.

Participants were free to choose the gesture to perform for each verb. The verbs used were selected to ensure participants would be able to perform a gesture with their upper body and face. Participants first completed a practise trial consisting of two verbs, to determine that they understood the instructions, and then completed the study phase. I sat behind the participant during the study and recorded their responses to ensure they were performing the task as instructed. During the study phase, verbs were visible for 3 seconds with a 1 second interval between verbs. This duration was tested before the study, to create an optimal time for the participant to execute a gesture, without the gesture becoming unnaturally long. The total duration of this phase was approximately 10 minutes.

Interference task. After the study phase participants completed a cognitive interference task to clear short-term memory/working memory to minimise the possibility that participants would adopt strategies to remember verbs (e.g., reciting verbs) and to minimise recency effects. Participants were asked to recite the alphabet backwards out loud. If a mistake was made, they were told to restart from the letter ‘z’. The duration of this task was approximately 3-5 minutes.

Free recall test. Following the interference task, I asked participants to freely recall as many verbs from the study list as they could. Participants wrote any verbs they could remember on a piece of paper without a time limit. I instructed participants to recall verbs regardless of colour, order, or mode of production. They were also made aware that guesses were acceptable to write down, and they would not be penalised for incorrect answers. All participants completed the recall phase within 15 minutes.

Recognition test. Participants then returned to the computer to perform an untimed recognition task. 120 verbs appeared on the screen one at a time. Two thirds of the verbs presented were the original study list (40 produced; 40 read) and the remaining third were 40 new verbs that were not previously studied. This phase took approximately 20 minutes to complete. When a verb was presented, participants were asked to press '/' (right hand) if they thought the verb was from the study list, and 'z' (left hand) if they thought it was a new verb (i.e., not one from the study list).

Source memory task. When participants indicated a verb was from the study list (i.e., they recognised it), they were asked to make a 'remember-know' judgment (Fawcett & Ozubko, 2016), to determine how well they remembered the verb. For these remember-know judgements, if the participant recollected both the verb and details of the context which they learned it, such as the colour of the letters, whether they said it aloud, or whether they verb was gestured or read silently, they were asked to respond that they 'know' the verb by pressing '/' (right hand). This indicated that the verb was represented strongly in memory. Alternatively, if they felt the verb was only familiar to them, and they did not know any of the context information of when the verb was learned (i.e., they only recall it as a verb that was presented), they were asked to respond that they 'remembered' the verb, by pressing 'z'. This represented a relatively weaker memory.

Debrief. Once the remember-know task was completed participants were de-briefed about the study's aims. Following this, participants were thanked for their participation and the study was completed.

Design

A quantitative, mixed subjects design was conducted. The independent variables were Group (between subjects) with two levels (gesture, and gesture/spoken), and Word Production (within subjects) with two levels (produced, silent). The dependent variables measured were free word recall, word recognition, and remember-know judgment information.

Results

Data was screened for errors prior to analysis. No errors were identified, and all participants were included in the final data set that was analysed. Verb recall distribution was mapped in order of study, and there appeared to be no patterns within the recall data, suggesting little to no recency or primacy effects, or ordering effects. Data was normally distributed, with no significant outliers. All descriptive statistics have been transformed from rates into percentages, for ease of interpretation.

Free Recall

A mixed factorial 2x2 ANOVA was conducted on the proportions of free recall hit rates (Figure 1). This tested the impact of enactment and speaking on recall memory, to determine whether established memory traces exist due to enactment and production. It was revealed there was a significant main effect of word type, F (1, 10) = 44.83, p < .001, d = 2.87, such that participants freely recalled more produced verbs (M = 39.0%, SD = 16.4, 95% CI [27.9, 50.0]) than read verbs (M = 6.3%, SD = 3.1, 95% CI [4.2, 8.3]). There was no significant main effect of production group, F (1, 10) = 0.04, p = .85, d = 0.11, such that the gesture/spoken group (M = 22.1%, SD = 9.0, 95% CI [13.9, 30.3]) and the gesture only group (M = 23.1%, SD = 9.0, 95% CI [14.9, 31.3]), did not differ in verb recall rates. There was no significant interaction effect involving production group and word type, F (1, 10) = 0.02, p = .90. The difference between groups for produced verbs and read verbs and silently read verbs were small, and not significant (Table 1).

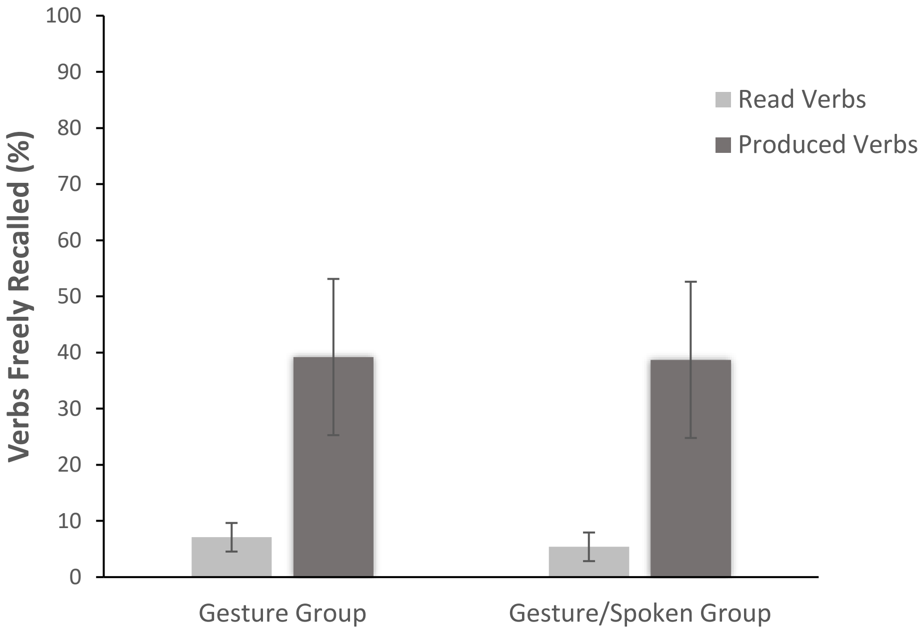


Figure 1: Mean percentage rates of correctly identified verb recall for produced and read verbs, separated by the different production condition groups. Error bars show 95% confidence intervals.

Table 1

Estimated Marginal Means for Interaction Effect on Recall Data

Group	Gesture		Gesture/Spoken	
	Mean	SD	Mean	SD
Word Type				
Read	7.1	3.7	5.4	2.5
Produced	39.2	16.2	38.7	18.1

Note. Output has been transformed from ratios to percentages for ease of interpretation. SD = standard deviation.

Recognition

A mixed factorial 2x3 ANOVA was conducted on the recognition rates for correct identification of verbs as being 'old' or 'new' (Figure 2). There was a significant main effect of word type, F (2, 20) = 9.54, p = .001. Post hoc analysis revealed that participants correctly recognised more produced verbs (M = 96.9%, SD = 4.5, 95% CI [93.8, 99.9]) than read verbs (M = 59.6%, SD = 20.2, 95% CI [46.1, 73.1]), t (11) = 6.27, p < .001, 95% CI [24.2, 50.4], d = 2.54. They also correctly identified more produced verbs than new verb correct rejection (M = 69.9%, SD = 22.4, 95% CI [55.0, 84.9]), t (11) = 4.03, p = .002, 95% CI [12.3, 41.6], d = 1.69. There was no significant difference between correct rejection of new verbs and read verb hits, t (11) = 0.89, p = .40, 95% CI [-15.4, 36.1], d = 0.48. There was no significant main effect of production group, F (1, 10) = 0.001, p = .97, d = 0.22, and therefore no overall difference between gesture group (M = 75.4%, SD = 4.0, 95% CI [71.1, 79.8]), and gesture/spoken group (M = 75.5%, SD = 5.0, 95% CI [71.2, 79.9]). There was no significant interaction between production group and word type, F (2, 20) = 0.22, p = .80. The difference between groups for produced verbs, read verbs, and silently read verbs were small, and not significant (Table 2).

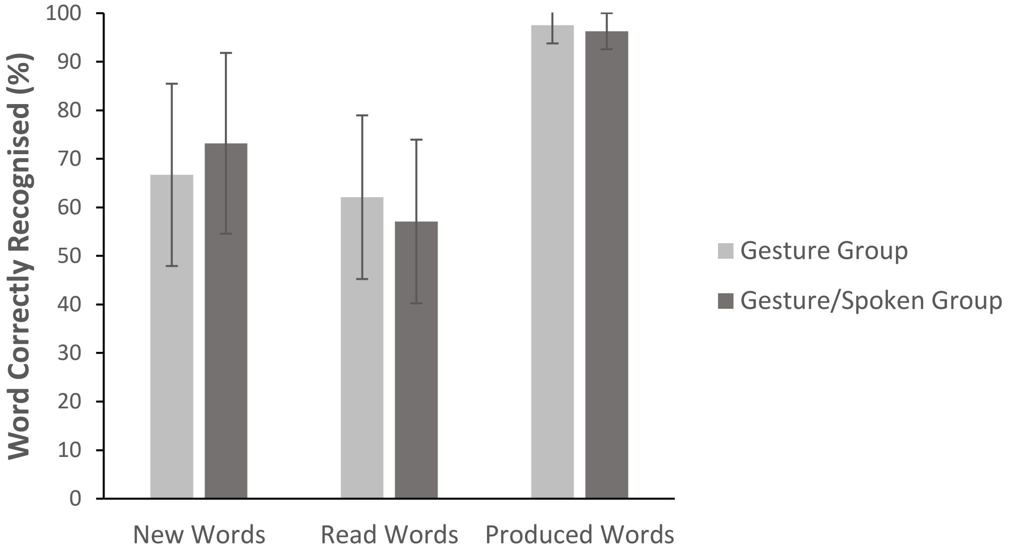


Figure 2: Percentage rates of correctly identified new, read and produced verbs as being ‘new’ or ‘old’, separated by the different production condition groups. Error bars show 95% confidence intervals.

Table 2

Estimated Marginal Means for Interaction Effect on Recognition Data

Group	Gesture		Gesture/Spoken	
	Mean	SD	Mean	SD
Word Type				
New	66.7	19.8	73.2	26.2
Read	62.1	18.4	57.1	23.3
Produced	97.5	3.2	96.3	5.9

Note. Output has been transformed for ratios to percentages for ease of interpretation. SD = standard deviation.

Remember-Know Judgments for Correctly Recognition

A paired-samples t-test was conducted on the remember-know judgments data, ignoring group type, to determine whether there were differences in produced and read verb 'know' judgments. These judgments were made by participants to indicate whether the verbs they recognised included recollection about information about the context in which the verb was studied (known), or if the verbs recognised were only familiar to the participant (remember). Enacted verbs had more 'know' responses, compared to verbs silently read, t (11) = 5.43, p < .001, 95% CI [21.9, 51.8], d = 1.87. Of the 59.6% of read verbs correctly identified, participants self-reported that 54.8% (SD = 27.0) were 'known' and 45.1% (SD = 27.0) were only 'remembered'. Of the 96.9% of produced verbs correctly identified, participants self-reported that nearly all, 91.7% (SD = 6.3), were confidently 'known' and only 8.3% (SD = 6.3) were 'remembered'. Participants were more likely to recognise and identify produced verbs as being words studied via enactment, than read verbs being studied via silent reading.

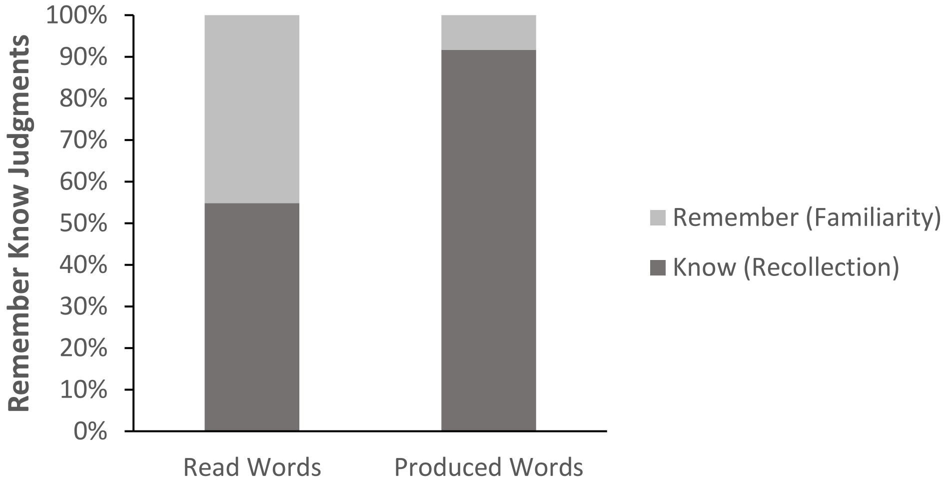


Figure 3: The percentage rates of remember-know judgments for both read and produced verbs that were correctly recognised.

Sensitivity of Recognition

As the recognition phase was 66.6% study trials and 33.3% new trials, it was important to determine whether participants showed sensitivity in their recognition. I did this via the use of a d' test (d' = Z [hit rate]/ Z [false positive rate]), which tests sensitivity while considering false alarm rates. This test helps identify if people are biased in their ability to recognise verbs as being previously studied. If participants were biased to make this decision, hit rates alone do not account for the large false alarm rates, and is not a good indicator for sensitivity. I determined that overall verb responses led to d' = 1.51 (*SD* = 0.51, 95% CI [1.19, 1.83]), suggesting that participants were highly sensitive in their ability to discriminate, and correctly recognise and identify verbs as being 'old' or previously studied. A mixed factorial 2x2 ANOVA was conducted on the d' recognition data to determine whether participants differed in sensitivity between groups and word types. It was revealed there was a significant main effect of word type, F (1, 10) = 55.22, $p < .001$, d = 2.43, such that participants were more sensitive to produced verbs (*M* = 3.03, *SD* = 1.13, 95% CI [2.27, 3.78]) than read verbs (*M* = 0.96, *SD* = 0.42, 95% CI [0.68, 1.23]). There was no significant main effect of production group, F (1, 10) = 0.22, p = .65, d = 0.27, such that the gesture/spoken group (*M* = 2.09, *SD* = 0.74, 95% CI [1.42, 2.76]) and the gesture only group (*M* = 1.89, *SD* = 0.74, 95% CI [1.22, 2.57]), did not differ in sensitivity rates. There was no significant interaction effect involving production group and word type, F (1, 10) = 0.02, p = .97. The difference between groups for produced verbs and silently read verbs were small, and not significant (Table 3).

Table 3

Estimated Marginal Means for Interaction Effect on d' Data

Group	Gesture		Gesture/Spoken	
	Mean	SD	Mean	SD
Word Type				
Read	0.86	0.34	1.05	0.50
Produced	2.92	1.09	3.13	1.26

Note. SD = standard deviation

Discussion

The current study addressed whether gesturally producing verbs led to greater recall and recognition rates, than silently reading verbs. It also addressed whether saying verbs aloud while gesturing them would lead to better recall and recognition for produced verbs, than if only gesturing them. As was expected, I was able to imitate comparable previous findings, and found an overall enactment effect in a within-subject, mixed-list design, in which unique distinctive responses were made to each item at study. There was no significant interaction effect for either recognition or recall data, suggesting that there was no additive effect of spoken production alongside the enactment effect on memory. Enactment led to greater 'know' judgments, suggesting that participants were more likely to identify that they 'knew' the context in which they studied verbs that they enacted, than read.

Extremely Large Enactment Effects

As was hypothesised, participants freely recalled more verbs they produced using gestures than verbs they read silently. Overall, participants correctly recalled

18 verbs on average, which was 20 – 25% of verbs presented, 87% of which were enacted. Of the 40 verbs produced, participants correctly recalled 38 – 40%. Participants only recalled an average of 2-3 of the 40 read verbs. This produced an enormously large effect (d = 2.87), suggesting an overwhelming enactment effect was present. As was also hypothesised, participants correctly recognised more verbs that they gestured than verbs they read. This again suggests that gesturing has a very large effect on memory (d = 2.54). Participants overall correctly identified 75-80% verbs as being ‘old and ‘new’. Enactment increased memory recognition rates from 60% to 96 - 98%.

As was expected, participants were overall better recognising verbs as being studied, than they were at free recalling studied verbs. Despite this, there was a larger overall enactment effect on recall memory (d = 2.87) compared to the already very large recognition memory enactment effect (d = 2.54). This signifies that enacting may impact free recall, which relies on the independent recollection of verbs’ cues to activate appropriate memory traces, more than recognition memory, where related cues are already provided at test to cause recollection or familiarity.

No Additive Effect of Verbal Production on the Enactment Effect

The second hypothesis, that speaking while gesturing would lead to greater recall and recognition than only gesturing, was not supported with the results of this study. Regarding the recognition data, participants recognised 39 of the 40 verbs that were produced at test. This led to significant ceiling effects, because of how overwhelmingly effective the overall enactment effect was on verb recognition. Due to this, I am unable to make definitive claims about the additive effect of speaking alongside enacting on recognition memory.

There was no interaction or main effect found within the recall data. I was expecting an interaction effect rather than a main effect of group, as I proposed that the addition of spoken production would increase ‘produced’ verb recall, while not impacting silent verb recall. Despite this, it seems adding vocal production alongside enacted production does not positively or negatively impact either produced verb or silent verb recall, showing that while speaking alongside gesturing does not seem to increase produced word recall memory, it also does not decrease it. The vocal production effect could not be found when the main encoding task (enacting) had already encouraged excellent memory.

Enactment Affects Judgments Around Source Memory

Overall there was a higher ratio of ‘know’ judgments for produced verb than silent verb recognition. This is likely due to an increase in distinctiveness, which solidified memory traces, allowing greater access to surrounding context information. As recognition tests all memory traces, those that were weak likely to not lead to correct recognition, moderate traces likely led to recognition, but only through familiarity, and strong traces likely led to activation of not only the verb, but memories of source information (indicated with a ‘know’ judgement).

Both theory and evidence (previous and current) seem to support that using semantically related production of words improves word memory; and, that deeper semantic and unique detailing in the expression of the word increases the cognitive encoding and distinctiveness of the word itself in comparison to words read silently. This suggests that producing words as gestures would lead to greater memory, than silently reading words. This increase in word memory can be identified in this study through the increased ‘know’ judgments for enacted words. Enactment allows memory of words to be remembered via recollection, rather than only familiarity.

However, recognition of read words relies more heavily on familiarity alone, rather than the more significant memories, such as recollection of study context and distinctive traits.

Overall Enactment Effect in Relation to Theory

This study's results support the distinctiveness account, which proposes that remembering items involves the activation of distinctive memory traces created at study. The more solid these distinctive traces are encoded; the more likely items will be remembered. When recalling a word, people are only able to independently tap into the most solid traces that were created (e.g., the distinctive produced words). However, when recognising a word all traces are tested, even those that are less solid (e.g. some read word traces). Recognition tasks test less solid traces with strong activation, which increases the chance that read words (which the majority are expected to be weak traces), as well as enacted words (which are expected to be at least moderately solid traces), are likely to be recognised. However, due to the limited number of participants, it may be too premature to make definitive claims around this.

Results suggest that while enactment increased enacted verb recognition, it may not have been at the expense of read verb memory. There were no significant differences in read verb recognition rates and the rates in which participants correctly recognised new verbs as being new unstudied verbs. As recognition rates for read verbs were comparable to that of the rates of correct filler verb identification as being new unstudied verbs, this can be used as a reference point to a baseline of memory accuracy.

Why No Additive Effect was Found Regarding Theory

MacLeod, et al. (2010) found that adding verbal production alongside the generation effect led to a 15-22% increase to hit rate. They conducted a simple within subject design, where participants generated all words from a cue. However, half of responses were spoken, and the others silently generated. This differed to the current studies design, which looked at the addition of speaking in a mixed factorial design, meaning we analysed the difference of verbal producing between groups, without sacrificing distinctiveness. While this study's analysis of verbal production as an additive effect was between group, each group studied a mixed list (a list where half of the words were simply read), to ensure distinctive traces would still be created for enacted (and produced) items.

It is possible that the design differences between MacLeod, et al.'s (2010) within-subject study and this between-groups study analysis, may have been what led to different results regarding the addition of vocal production to another prominent memory effect. However, I suggest there are two alternative explanations as to why this study did not find the hypothesised results. Firstly, the production effect has not been found to impact recall rates, and therefore additive effects would not be expected to be found for this study's recall rates. Secondly, ceiling effect prevented any interpretation of whether the vocal production effect existed alongside the enactment effect in recognition rates.

The increased effect on memory MacLeod, et al., (2010) found due to vocal production, when combined with the generation effect, was equivalent to the traditional production effects size. While this studies design differed, I had also expected an increase of recognition accuracy, and potentially recall rates, of similar size as the verbal production effect (10-20%). However, the verbal production effect has only ever been consistently found to impact recognition rates, and not recall rates (MacLeod, et al., 2010). This could explain why no additive effect of verbal production was found with the recall data. I am also unable to determine whether verbal production affected recognition accuracy due to ceiling effects. While I cannot make any conclusive statements with the results of this study regarding the second hypothesis, it is also possible that the enactment effect may have just overwhelmed the production effect.

Does the Enactment Effect Occur Due to Distinctiveness or Strength?

The list-strength theory (Ratcliff, et al., 1990; Verde, 2009), refers to the phenomena where memory benefits of producing a mixed-list, due to increased produced word strength, are often to the detriment of read word memory. It has been that the production effect is not impacted by the list-strength effect, as a significant decrease in silent items memory has not been observed in a mixed list compared to a pure list design (MacLeod, et al., 2010). It is said that verbal production increases distinctiveness of items, rather than their encoded strength, which increases their ability to be correctly recognised as being previously studied (MacLeod, et al., 2010). It is because of this that the list-strength effect, which only impacts strength, does not seem to impact the vocal production effect. It has also been suggested that as the list-strength effect does not occur in recognition, recognition must rely on distinctiveness rather than strength (MacLeod, et al., 2010). Therefore, as strength does not seem to have any bearing on recognition, and the production effect is observed in recognition, it is suggested that the production effect does not rely on strength.

Unlike the production effect, it seems the enactment effect does not only influence recognition rates, but also recall rates, which has been known to be impacted by the list-strength effect. I therefore suggest that the enactment effect has a more complicated relationship to strength and distinctiveness than the production effect, which overwhelming seems to be determined by distinctiveness. The enactment effect likely depends on a combination of both strength and distinctiveness.

Recognition is suggested to be facilitated by distinctiveness rather than strength. Initial analysis comparing correct rejection of new verbs to the hit rates of read verbs suggests that the list-strength effect does not seem to impact this experiment's recognition rates. If this is true, the enactment effect on recognition is likely to be due to verb distinctiveness.

The list-strength effect does not normally affect recognition, only recall. I suggest that this is because only recall, and not recognition, relies on word strength. Logically, this makes sense. Only the strongest encoded verbs would be able to be freely recalled from memory without a cue. However, recognition relies more on the discrimination of stimuli, and distinctiveness would be more useful in effectively doing this.

Results found that enactment significantly impacted recognition, suggesting that discrimination play an important role in the enactment effect. However, the enactment effect was also observed to impact recall data, which may suggest that enactment also increases word strength. Even though definitive claims cannot be made around the list-strength effect existing in this study's recall data, if it was assumed there was a list-strength effect, then the enactment effect on recall is likely due to strength.

As previously discussed in the introduction, it is suggested that 'strength' plays an important role in how distinctive traces are embedded in memory. The strength of the items enacted may strengthen the distinctive traces created by enacting and embodying the verbs meaning, which improves recall. For example, larger gross movements (e.g. punching, flying) might produce a stronger consolidation of the distinctiveness trace than smaller fine movements (e.g. breathing, blinking), because requires flying more creative, cognitive, and bodily effort, which increases the strength of the distinctive traces created by producing. A way to potentially test this is to enforce a high level of encoding for all verbs. This could be done by asking participants to make a semantic judgement for each verb to force semantic analysis of all verbs and strengthen encoding. This should eliminate any benefit enacting has on recall due to additional strength of half of the list items.

Limitations

Small sample size. There are a several limitations to this study. Firstly, due to set backs with ethics approval, there was a severe delay in participant recruitment which led to a decrease in overall participant numbers. A conservative a-priori power analysis determined 40 participants would be desirable for this investigation, unfortunately I only collected 12 participants. However, despite this I found very large effects for the within subject memory rates. This was expected, as large effects have been found before with comparable enactment designs. However, I may have not collected enough participants to find an overall smaller interaction effect throughout all the individual difference between participants in each group, with the mixed design chosen for this experiment. Furthermore, due to the small sample size, this also limits the results applicability and how they can be generalised to the student population.

Ceiling effects. It appears the enactment effect was so effective in this study, as participants recognised almost all 40 enacted verbs at test, that I was not able to determine if there was an additive effect with vocal production. Previous studies of the production and enactment effects have also only used eighty-verbs total in their learning task, which made these ceiling effects unexpected. The number of verbs for the study and test phases were decided on after review of previous enactment studies. Some studies suggested that without contextual information in the form of phrases, or props to act onto, verbs would not be as effectively remembered (Steffens, et al., 2003). However, this certainly does not seem to be the case here.

I suggest there are two reasons for such a large enactment effect being found. First, participants only had to recall and recognise single verbs, rather than larger chunks of information (e.g. action phrases) which involve more complex information to encode. Secondly, as objects were not explicitly stated alongside the verbs, as was the case in previous enactment studies which used verb phrases, the lack of objects or contextual props in this study did not impact participants' ability to remember the single verbs. Future studies should therefore add to the comprehensive list of verbs created for this study or create their own larger list of words, study phrases, or material of their choice.

Memory test ordering. A limitation of the measures used in this study include the ordering of each test phase. As recall was measured before recognition, this may have led to additional elaboration and re-ignition of memory traces for both the items correctly freely recalled, but also may have led to additional activation of associated memory traces close to the other studied verbs, potentially making them easier to recognise due to partial activation. However, it was necessary that recall was measured before recognition, as measuring recognition first would have created a worse impact on participants than re-studying the original list, as they would be re-exposed to the original verbs a second time, reactivating and consolidating their memories further. Future studies could address this by only focusing on measuring one type of memory, or perhaps adding a secondary filler tasks between test conditions.

Directions for Future Research

Addressing the list-strength theory. It is important to consider the list-strength theory (Verde, 2009) regarding the limitations of this study, which is the finding that increasing the strength of some items in a word list would decrease the strength of other list items (Ratcliff, et al., 1990). Ideally, studies want to limit the list-strength effect as much as possible, so the memory benefits of enacting or verbally producing do not negatively impact rates in which silently read verbs are remembered. This study did not gather a baseline to measure the comparative difference in memory rates if the study list was purely learnt via only weak (reading silently) or only strong (enacting and/or speaking) means. If that data was collected I would have been able to determine whether read verb recall and recognition rates were significantly worse for mixed-list design, compared to a pure-list design.

Future studies may wish to gather baseline information to determine whether the list-strength effect impacts recognition and recall memory rates in this mixed-list design. Especially since the list-strength effect is often very prominent in free recall paradigms, however, is quite absent in recognition (Verde, 2009). As previously mentioned, while I did not collect the data to address whether the strength of the enacted verbs damaged the memory strength of the read verb memories, by eyeballing the data, it seems like a likely possibility this this was the case with the recall data, but not with the recognition data.

Within-subject vs between-subject design. Future studies should opt for a completely within subject design, comparing gesturing to gesture/speaking within participants, to determine whether adding speaking to the enactment effect would increase verb memory. Using a similar design to that of MacLeod, et al.'s (2010) generation and production effect study may lead to clearer results that are more likely to pick up on a production effect if it exists alongside the enactment effect. They should also use larger list of study verbs, to avoid recognition ceiling effect.

Long-term memory benefits. In this study the time between the study and test phases was quite brief; approximately 15 minutes to complete the filler task. This means the memories being tapped into were still very 'fresh' and recent activated in memory and not necessarily completely consolidated in long-term memory. Future research could test this enactment effect on long-term memory to determine whether these benefits persist over time. The use of multiple session of study would also be useful when investigating long term benefits, as it would be useful to see the exact benefit of each additional study session.

Alternative combinations of memory effects. Future research could look at the benefit of adding alternative effects together to determine which combination of memory effects would be most effective with different studying material. The addition of props to use during enactment in comparison to simply miming gestures onto imagined objects, could also be an option for future research to investigate. Future studies wishing to replicate these findings may want to include a separate baseline, to determine whether new verb and read verb correct recognition rates are both or neither impacted by the list-strength effect (Verde, 2009), in a mixed list enactment experiment.

Use of full body in enactment. This study specifically investigated the effect of upper body enactment of verbs. Further investigation could address the use of entire body movement, and the extent of which larger movements impact memory in comparison to smaller movements. As vocal production has been shown to be more effective with more ‘force’ used, with singing being more effective than loudly saying aloud, which is better than whispering, which is better than mouthing, which is better than reading silently (MacLeod, et al., 2010; Quinlan, & Taylor, 2013). Perhaps this effect could be found with more elaborate and purposeful movements such as full body enactment.

Different population samples. This study looked at a predominantly student population, which was important as the justification for this research was to help determine whether enactment could be used as an effective studying tool. However, future research may want to use a more representative population, or other populations that may specifically benefit from the enactment effect. For example, future studies could investigate these effects regarding children, who have low levels of prior learning, and learning rates are more accelerated than adults. Alternatively, future studies could investigate older adults, those suffering with Alzheimer’s and dementia, people diagnosed with a traumatic brain injury, or people with specific language-based deficits. Research into facilitating these populations memory retention is meaningful and important in developing strategies to help these populations ease of living.

Alternative study material. While this study is a good foundation to enactment effect research, it was quite limited in the study material, only using verbs that can be performed using the upper body and face. It would be interesting to investigate different study material, perhaps the enactment of nouns, e.g. ‘cat’ could be enacted by a person acting like a cat. While the production effect has looked at its effectiveness of learning studying material, it would be useful to see whether the enactment also could be used to study more real-world information.

Increasing second language acquisition. While only single verbs were studied in this experiment, previous research has also looked at the enactment of phrases, finding an overwhelming enactment effect (Engelkamp, et al., 1995). Learning verb-related words and phrases via embodiment, while learning words that cannot as easily be enacted (e.g., nouns and adjectives) by reading them silently, could lead to faster acquisition of a second language, compared to learning words all silently or vocally. These findings lend themselves as a basis of future research that could explore the possibility of using embodiment to perhaps facilitate second language acquisition.

Implications and Application

It is important to note that while study did not find evidence to support that combined vocal and gestured production have any cognitive benefit to the person attending to the information, gesturing could be used as a tool to facilitate communication between those with traditional language processing deficits. This would be to the benefit of both the communicator and the listener, allowing for easier memorising of information communicated between the two parties. The communicator could both vocalise and gesture instructions, and the listener could repeat back the gesture motions, to signal understanding, and consolidate the memories of the newly learnt instructions.

The results support the use of embodiment in learning written material. Traditionally bodily movement has been used as a tool to learn through means of modelling and copying someone else completing the action (e.g. tying shoelaces) or used in personal repetition to learn new physical skills (e.g. writing). However, schools and other learning environments may benefit from adopting the use of enactment and bodily production to learn and remember written information.

Students could embody concepts to facilitate learning, and they could still be tested with a traditional MCQ, which relies on recognition. While this study suggests that enacted information can be tested using written/typed means, Objective Structured Clinical Exams (OSCE's) are still an option for testing material studied via enactment, as the material being remembered during examination would be further reinforced through practice.

Summary and Conclusion

This study proposes that there are no additive benefits of vocalising verbs when performing them as gesture during study. However, there are extremely large within subject effects, suggesting mimed performance of verbs significantly increases memory. This effect was found with both recall and recognition memory performance, although recognition led to better overall memory accuracy than free recall. Enacting verbs may also increase the amount contextual meta-cognitive contextual information recalled during recognition. An enormous effect for producing verb items as mimes over silent reading can clearly be observed. This research gives evidence for incorporating a more hands on approach to studying in several environments. This can include: incorporating gesturing in second-language learning, greater emphasis on OSCE's in educational settings, or even adopting interpretive dance as an alternative studying method. Due to some of the unforeseen limitations of this study, further investigation is necessary to address whether using multiple learning strategies can have an additive effect on studying.

References

Bodner, G.E., Jamieson, R.K., Cormack, D. T., McDonald, D. L., & Bernstein, D. M. (2016). The production effect in recognition memory: Weakening strength can strengthen distinctiveness. *Canadian Journal of Experimental Psychology, 70*(2), 93-98. Doi: 10.1037/cep0000082.

Bodner, G. E., Taikh, A., & Fawcett, J. M. (2014). Assessing the costs and benefits of production in recognition. *Psychonomic Bulletin & Review, 21,* 149-154. doi: 10.3758/s13423-013-0485-1.

Castel, A. D., Rhodes, M. G., & Friedman, M. C. (2013). Predicting memory benefits in the production effect: The use and misuse of self-generated distinctive cues when making judgements of learning. *Memory & Cognition, 41,* 28-35. doi: 10.3758/s13421-012-0249-6.

Conway, M. A., & Gathercole, S. E. (1987). Modality and long-term memory. *Journal of Memory and Language, 26,* 341–361. doi: 10.1016/0749-596X(87)90118-5.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11,* 671–684. doi: 10.1016/S0022-5371(72)80001-X.

Dewhurst, S. D., Rackie, J. M., & van Esch, L. (2016). Not lost in translation: Writing auditorily presented words at study increase correct recognition "at no cost*". Journal of Cognitive Psychology, 28*(4), 437-442. doi: 10.1080/20445911.2016.1145684.

Engelkamp, J., Zimmer, H. D., & Kurbjuweit, A. (1995). Verb frequency and enactment in implicit and explicit memory. *Psychological Research, 57,* 242-249. doi: 10.1007/BF00431285.

Fawcett, J. M. (2013). The production effect benefits performance in between-subject designs: A meta-analysis. *Acta Psychologica, 142*, 1-5. doi: 10.1016/j.actpsy.2012.10.001.

Fawcett, J. M., Quinlan, C. K., & Taylor, T. L. (2012). Interplay of the production and picture superiority effect: A signal detection analysis. *Memory, 20*(7), 655-666. Doi: 10.1080/09658211.2012.693510.

Fawcett, J. M., & Ozubko, J. D. (2016). Familiarity, but not recollection, support the between-subject production effect in recollection memory. *Canadian Journal of Experimental Psychology, 70*(2), 99-115. doi: 1037/cep0000089.

Forrin, N. D., Jonker, T. R., & MacLeod, C. M. (2014). Production improves equivalently following elaborative vs non-elaborative processing. *Memory, 22*(5), 470-480. doi: 10.1080/09658211.2013.798417.

Forrin, N. D., & MacLeod, C. M. (2017). This time it's personal: The memory benefit of hearing oneself. *Memory, 26*(4). 574-579. doi: 10.1080/09658211.2017.1383434.

Forrin, N. D., Macleod, C. M., & Ozubko, J. D. (2012). Widening boundaries of the production effect. *Memory & Cognition, 40*, 1046-1055. doi: 10.3758/s13421-012-0210-8.

Gathercole, S. E., & Conway, M. A. (1988). Exploring long-term modality effects: Vocalization leads to best retention. Memory & Cognition, 16, 110–119. doi: 10.3758/BF03213478.

Hopkins, R. H., & Edwards, R. E. (1972). Pronunciation effects in recognition memory. *Journal of Verbal Learning and Verbal Behavior, 11*(4), 534–537. doi: 10.1037/h0036825.

Hourihan, K. L., & Smith, A. R. (2016). Production does not improve memory for face-name association. *Canadian Journal of Experimental Psychology, 70*(2), 147-153. doi: 10.1037/cep0000091.

Icht, M., & Mama, Y. (2015). The production effect in memory: A prominent mnemonic in children. Journal of Child Language, 42(5), 1102-1124. doi: 10.1017/s0305000914000713.

Jones, A. C. & Pyc, M. A. (2013). The production effect: Costs and benefits in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* Advance online publication. doi: 10.1037/a0033337.

Lin, O. Y. H., & MacLeod, C. M. (2012). Aging and the production effect: A test of the distinctiveness account. *Canadian Journal of Experimental Psychology, 66*(3), 212-216. doi: 10.1037/a0028309.

MacLeod, C. M. (2011). I said, you said: The production effect gets personal. *Psychonomic Bulletin & Review, 18,* 1197-1202. doi: 10.3758/s13423-011-0168-8.

MacLeod, C. M., Gopie, N., Hourihan, K. L., Neary, K. R., & Ozubko, J. D. (2010). The production effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(3), 671-685. doi: 10.1037/a0018785

Ozubko J. D., Gopie N., MacLeod C. M. (2012). Production benefits both recollection and familiarity. *Memory & Cognition, 40,* 326–338. doi: 10.3758/s13421-011-0165-1.

Ozubko, J. D., Hourihan, K. L., & MacLeod, C. M. (2012). Production benefits learning: The production effect endures and improves memory for text. *Memory, 20*(7), 717-727. doi: 10.1080/09658211.2012.699070.

Ozubko, J. D., & MacLeod, C. M. (2010). The production effect in memory: Evidence that distinctiveness underlies the benefit. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(6), 1543- 1547. doi: 10.1037/ a0020604.

Ozubko, J. D., Major, J., & MacLeod, C. M. (2014). Remembered study mode: Support for the distinctiveness account of the production effect. *Memory, 22*(5) 509-524. doi: 10.1080/09658211.2013.800554

Peirce, J.W. (2009). Generating stimuli for neuroscience using PsychoPy. *Front. Neuroinform. 2*(10). doi: 10.3389/neuro.11.010.2008.

Quinlan, C. K., & Taylor. T. L. (2013). Enhancing the production effect in memory. *Memory, 21*(8), 904-915. doi: 10.1080/09658211.2013.766754

Ratcliff, R., Clark, S. E., & Shiffrin, R. M. (1990). List-strength effect: I. Data and discussion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(2), 163–178. doi: 10.1037/0278-7393.16.2.163.

Sidhu, D. M., Kwan, R., Pexman, P. M., and Siakaluk, P. D. (2014). Effects of relative embodiment in lexical and semantic processing of verbs. *Acta Psychol. 149*, 32–39. doi: 10.1016/j.actpsy.2014.02.009.

Sidhu, D. M., & Pexman, P. M. (2016). Is moving more memorable than proving? Effects of embodiment and imagined enactment of verb memory. *Frontiers in Psychology, 7*(1010), 1-13. doi: 10.3389/fpsyg.2016.01010.

Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. Journal of Experimental Psychology, *Human Learning and Memory, 4*(6), 592–604. doi: 10.1037/0278-7393.4.6.592.

Steffens, M. C., Buchner, A., & Wender, K. F. (2003). Quite ordinary retrieval cues may determine free recall of actions. *Journal of Memory & Language, 48,* 399-415.

Steffens, M. C., Buchner, A., Wender, K., F., & Decker, C. (2007). Limits on the role of retrieval cues in memory for actions: Enactment effect in the absence of object cues in the environment. *Memory & Cognition, 35*(8), 1841-1853. doi: 10.3758/BF03192919.

Steffens, M. C., von Stulpnagel, R., & Schult, J. C. (2015). Memory recall after “learning by doing” and “learning by viewing”: Boundary condition of an enactment effect. *Frontiers in Psychology, 6*(1907), 1-10. doi: 10.3389/fpsyg.2015.01907.

Taikh, A., & Bodner, G. E. (2016). Evaluating the basis of the between-group production effect in recognition. *Canadian Journal of Experimental Psychology, 70*(2), 186-194. doi: 10.1037/cep0000083.

Verde, M. F. (2009). The list-strength effect in recall: Relative-strength competition and retrieval inhibition may both contribute to forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(1), 205–220. doi: 10.1037/a0014275.

Wammes, J. D., Meade, M. E., & Fernandes, M. A. (2017). Learning terms and definition: Drawing and the role of elaborative encoding. *Acta Psychologica, 179*, 104-113. doi: 10.1016/j.actpsy.2017.07.008.

Wammes, J. D., Meade, M. E., & Fernandes, M. A. (2016). The drawing effect: Evidence for reliable and robust memory benefits in free recall. *The Quarterly Journal of Experimental Psychology, 69*(9), 1752-1776. doi: 10.1080/17470218.2015.1094494.

Appendix A

List of 80 Verbs Involving the Upper Body and Face Used in the Study Phase, including the 40 **enacted and/or spoken verbs**, ordered by presentation; top to bottom, left to right.

paddle	**pack**	inhale	**shake**
file	unbutton	stare	claw
catch	**salute**	swim	**wipe**
peel	tear	**call**	**carry**
swallow	**spray**	snatch	paint
knit	tickle	nod	**fold**
massage	sweep	**squeeze**	**tie**
milk	**comb**	**dig**	**spill**
blow	shrug	scare	**wave**
scrape	squash	drum	**tap**
poke	fasten	**uncap**	whisk
relax	strum	search	**fly**
draw	**stamp**	switch	stir
choke	**frown**	**scratch**	point
tremble	**cough**	drop	twist
hang	whip	**wriggle**	**ski**
give	punch	**curl**	**slap**
slouch	**cross**	trace	
fiddle	unscrew	knock	
tick	**cover**	**smoke**	
pinch	**pray**	**lick**	

Appendix B

List for 120 Verbs Involving the Upper Body and Face Used in the Recognition Test, including the 40 *newly presented verbs*, ordered by presentation; top to bottom, left to right.

whip	pray	ski	*bend*	knock	*sigh*
unscrew	spill	draw	pack	carry	*flex*
swallow	paint	*pull*	whisk	*flick*	sweep
hang	*press*	wriggle	*smile*	stir	*grab*
curl	poke	*saw*	*drive*	*scribble*	drum
open	*yawn*	pinch	*stitch*	give	*stack*
trace	*scrub*	knit	squash	*throw*	tie
snatch	slap	squeeze	spray	*type*	*shave*
lock	tremble	inhale	fold	scare	relax
extend	stare	cover	wipe	*knead*	cross
drop	*inject*	scrape	massage	*chop*	strum
dust	shrug	catch	point	swim	smoke
pat	*surprise*	milk	*clap*	dig	paddle
hug	salute	*kiss*	tickle	blow	*stretch*
call	claw	*hammer*	punch	choke	nod
cough	*drag*	*close*	*drink*	wave	*stroke*
comb	tap	unbutton	*push*	file	frown
fly	search	*juggle*	twist	stamp	peel
pour	uncap	lick	fiddle	slouch	scratch
tear	fasten	*clench*	switch	shake	tick

Appendix C

Ethics Approval Letter

Social Science Ethics Officer
Private Bag 01 Hobart
Tasmania 7001 Australia
Tel: (03) 6226 2763
Fax: (03) 6226 7148
Katherine.Shaw@utas.edu.au

HUMAN RESEARCH ETHICS COMMITTEE (TASMANIA) NETWORK

UTAS

16 August 2018

Dr Michael Garry
Psychology
Private Bag 30

Dear Dr Garry

Re: MINIMAL RISK ETHICS APPLICATION APPROVAL
Ethics Ref: H0017563 - Saying and Doing: Verbs and the Production Effect

We are pleased to advise that acting on a mandate from the Tasmania Social Sciences HREC, the Chair of the committee considered and approved the above project on 16 August 2018.

This approval constitutes ethical clearance by the Tasmania Social Sciences Human Research Ethics Committee. The decision and authority to commence the associated research may be dependent on factors beyond the remit of the ethics review process. For example, your research may need ethics clearance from other organisations or review by your research governance coordinator or Head of Department. It is your responsibility to find out if the approval of other bodies or authorities is required. It is recommended that the proposed research should not commence until you have satisfied these requirements.

Please note that this approval is for four years and is conditional upon receipt of an annual Progress Report. Ethics approval for this project will lapse if a Progress Report is not submitted.

The following conditions apply to this approval. Failure to abide by these conditions may result in suspension or discontinuation of approval.

1. It is the responsibility of the Chief Investigator to ensure that all investigators are aware of the terms of approval, to ensure the project is conducted as approved by the Ethics Committee, and to notify the Committee if any investigators are added to, or cease involvement with, the project.

2. Complaints: If any complaints are received or ethical issues arise during the course of the project, investigators should advise the Executive Officer of the Ethics Committee on 03 6226 7479 or human.ethics@utas.edu.au.

3. Incidents or adverse effects: Investigators should notify the Ethics Committee immediately of any serious or unexpected adverse effects on participants or unforeseen events affecting the ethical acceptability of the project.

4. Amendments to Project: Modifications to the project must not proceed until approval is obtained from the Ethics Committee. Please submit an Amendment Form (available on our website) to notify the Ethics Committee of the proposed modifications.

5. Annual Report: Continued approval for this project is dependent on the submission of a Progress Report by the anniversary date of your approval. You will be sent a courtesy reminder closer to this date. **Failure to submit a Progress Report will mean that ethics approval for this project will lapse.**

6. Final Report: A Final Report and a copy of any published material arising from the project, either in full or abstract, must be provided at the end of the project.

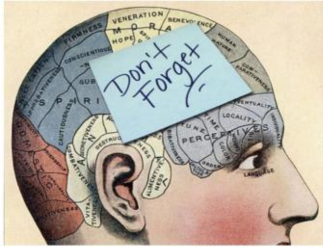
Yours sincerely

Jude Vienna-Hallam
Acting Executive Officer – Social Science
Tasmania Social Sciences HREC

Appendix D

Participant Recruitment Poster

UNIVERSITY of TASMANIA



VOLUNTEERS FOR MEMORY RESEARCH

Saying and Doing: Verbs and the Production Effect

We are looking for adult participants (aged 18 to 50 years old)

who speak English as a first language to take part

in one of three studies of word production and memory.

Your participation would take around **60-90 minutes.**
For participating, you will go into the draw to receive
one of **six $100 vouchers**.
(Psychology students will have the option to earn earn 60-90 minutes research participation credit instead.)

The study will involve you doing the following:

- Speaking, reading, or acting out words.
- Memorising and recollecting words, and
- Completing a computerised memory task.

For more information or to volunteer for this study:
Please contact Hayley Blowfield
hayleyb4@utas.edu.au
OR psychology students can apply through SONA

This study has been approved by the University of Tasmania Human Research Ethics Board. (H0017563)

Memory Study
Contact: Hayley Blowfield
Email: hayleyb4@utas.edu.au

Memory Study
Contact: Hayley Blowfield
Email: hayleyb4@utas.edu.au

Memory Study
Contact: Hayley Blowfield
Email: hayleyb4@utas.edu.au

Memory Study
Contact: Hayley Blowfield
Email: hayleyb4@utas.edu.au

Memory Study
Contact: Hayley Blowfield
Email: hayleyb4@utas.edu.au

Memory Study
Contact: Hayley Blowfield
Email: hayleyb4@utas.edu.au

Memory Study
Contact: Hayley Blowfield
Email: hayleyb4@utas.edu.au

Memory Study
Contact: Hayley Blowfield
Email: hayleyb4@utas.edu.au

Memory Study
Contact: Hayley Blowfield
Email: hayleyb4@utas.edu.au

Appendix E

Information Sheet

UNIVERSITY of TASMANIA

COLLEGE OF HEALTH and MEDICINE
School of Medicine

Participant Information Sheet version 1 16 July 2018

Saying and Doing: Verbs and the Production Effect

For participants

1. Invitation

You are invited to participate in a study investigating the production effect and its effect on memory. This study is part of psychology Honours research project of Hayley Blowfield, Michael O'Leary and Zac Bain-Williams. The study is supervised by Dr Michael Garry from the Division of Psychology at the University of Tasmania.

2. What is the purpose of this study?

The purpose of this study is to investigate whether the way people study words affects how well the studied words are remembered. The results from this study may help teachers design activities to make learning more effective. We can't provide too many details at the start of the study as this could affect the results. Once you have completed your study session we will fully explain the aims and hypotheses.

3. Why have I been invited to participate?

You have been invited to participate because you are between the age of 18 and 50, speak English as my first language, have normal or corrected to normal vision, are not colour blind, and don't have a reading disorder, or neuromuscular or neurological disorder.

4. What will I be asked to do?

The study will involve a single session of up to 90 minutes duration. The session will be in either the Psychology Research Centre, or the Social Sciences Building on the Sandy Bay campus at the University of Tasmania.

At beginning of the session you will be asked to provide some information about yourself, such as your age and sex, and will be asked to complete a questionnaire to determine your hand preference. The experimenter will then assign you to a group which will determine what you will be asked to do in the remainder of the session.

During the main part of the session you will be seated at a desk in front of a computer screen. The screen will present individual words, one after the other, several seconds apart. Some of the words will be written in white letters and other words will be written in blue letters. Approximately 80-100 words will be presented.

Regardless of which group you are in you will be asked to silently read all words written in WHITE letters. For words written in BLUE letters, what you will be asked to do will depend on which group you are in. You may be asked to say the blue words aloud, to perform a movement with your hands to mime the blue word, or to both say the word and perform the miming hand movement simultaneously. The experimenter will explain the precise instructions to you.

To allow us to confirm the accuracy of your vocal responses we will use a microphone to record your spoken responses. The recordings will be used to confirm that your response and how long it took you to respond. The recordings will be deleted once this data has been extracted.

5. Are there any possible benefits from participation in this study?

Your participation in this study will contribute to the scientific literature about a phenomenon called the production effect. This effect has been used to improve how students learn new information. This study will add to that body of knowledge.

Your participation in this study will provide valuable research training for three Psychology Honours students: Hayley Blowfield, Michael O'Leary and Zac Bain-Williams.

If you are a first year Psychology student you will receive up one hour course credit. If you are not a first-year psychology student, or aren't seeking course credit your name will be entered into a draw to win one of six $100 Coles-Myer vouchers.

6. Are there any possible risks from participation in this study?

There are no risks associated with participating in this study. If you become tired during the study, you can take an extended break between trials.

7. What if I change my mind during or after the study?

Participation in this study is entirely voluntary. You are welcome to withdraw from the study without consequence at any time if you wish. You are encouraged to do so if you are experiencing distress. If your data has already been collected you can request to have it removed and destroyed up to 14 days after the date of your participantion. It will not be possible to remove your data after this time as it will be de-identified for data analysis. However, once the data has been analysed and the thesis written your data will not be able to be removed.

8. What will happen to the information when this study is over?

All data will be replaced by an alphanumeric code, which will be recorded on your consent form. This will allow your data to be re-identified if necessary. Your identity will remain confidential for the purposes of the research and your information will not be provided to any sources without your knowledge. Any paper information will be stored of securely, in a locked cabinet at the University of Tasmania. Your electronic data will be stored securely on a password-protected computer at the University of Tasmania. Data will be kept for a minimum of 5 years after the date of publication.

9. How will the results of the study be published?

The final results will be reported an Honours thesis, which will be available to access through the UTAS Psychology Test Library. A brief summary of the thesis results will be available from 1st December 2018. Please contact the Chief Investigator, Dr Mike Garry (Michael.garry@utas.edu.au) if you would like to receive a copy of the summary.

10. What if I have questions about this study?

If you have any questions or concerns about the research, you are advised and encouraged to consult one of the student investigators (Hayley Blowfield, hayleyb4@utas.edu.au;

UNIVERSITY of TASMANIA

COLLEGE OF HEALTH and MEDICINE
School of Medicine

Participant Information Sheet version 1 16 July 2018

Michael O'Leary, olearym@utas.edu.au; Zac Bain-Williams, zbain@utas.edu.au) or Mike Garry (chief investigator) on michael.garry@utas.edu.au at any time.

This study has been approved by the Tasmanian Social Sciences Human Research Ethics Committee. If you have concerns or complaints about the conduct of this study, please contact the Executive Officer of the HREC (Tasmania) Network on +61 3 6226 6254 or email human.ethics@utas.edu.au. The Executive Officer is the person nominated to receive complaints from research participants. Please quote ethics reference number [H0017563].

Thank you for considering participating in our study.

You are welcome to keep this information sheet and refer back to it.

If you wish to participate in this study, please sign the attached consent form.

Appendix F

Consent Form

UNIVERSITY of TASMANIA

University of Tasmania

COLLEGE OF HEALTH and MEDICINE
School of Medicine

Participant Consent Form [Version 1] [16/07/2018]

Saying and Doing: Verbs and the production effect.

This form is for participants

1. I agree to take part in the research study named above.
2. I have read and understood the Information Sheet for this study.
3. The nature and possible effects of the study have been explained to me.
4. I understand that the study involves reading words presented on a computer monitor and that the study will take up to 90 minutes to complete.
5. I understand that a microphone will be used to record my voice during trials. This is to allow the researchers to confirm my responses have been accurately recorded. I also understand that these recordings will be erased once the accuracy of responses has been confirmed.
6. I understand that there are no risks associated with participating in this study
7. I understand that all research data will be securely stored on the University of Tasmania premises for a minimum of five years from the publication of the study results, and will then be destroyed.
8. Any questions that I have asked have been answered to my satisfaction.
9. I understand that the researcher(s) will maintain confidentiality and that any information I supply to the researcher(s) will be used only for the purposes of the research.
10. I understand that the results of the study will be published so that I cannot be identified as a participant.
11. I understand that my participation is voluntary and that I may withdraw at any time without any effect. If I so wish, I may request that any data I have supplied be withdrawn from the research , though this will only be possible up to 14 days following participation.

Participant's name: ____

Participant's signature: ____

Date: ______________________

UNIVERSITY of TASMANIA

COLLEGE OF HEALTH and MEDICINE
School of Medicine

University of Tasmania

Participant Consent Form [Version 1] [16/07/2018]

Statement by Investigator

☐ I have explained the project and the implications of participation in it to this volunteer and I believe that the consent is informed and that he/she understands the implications of participation.

If the Investigator has not had an opportunity to talk to participants prior to them participating, the following must be ticked.

☐ The participant has received the Information Sheet where my details have been provided so participants have had the opportunity to contact me prior to consenting to participate in this project.

Investigator's name: ____

Investigator's signature: ____

Date: ____________________

Appendix G

Debrief Statement

Production effect participant debrief statement.

Now that your data collection session is complete I would like to explain the aims and hypotheses.

The aim of this study is to investigate a research finding called the *production effect*. The production effect is the finding that when we try to remember a list of words, our memory is usually better for words that we have spoken out loud compared to words that we have only read silently.

In the study you just completed you spoke out loud, and/or mimed, some of the words that were presented, and silently read other words. Based on previous research on the production effect I expect to find that you will remember more words that you spoke out loud, or mimed, compared to words that you only read silently.

Another aim of this study is to investigate whether the production effect is found when the words studied are verbs (words that describe actions such as 'pinch', and 'hit'). This differs from previous studies which have used noun words (words that describe things such as apple, and car). One of the reasons I used verbs instead of nouns is because I want to know whether the production effect would also occur when people mime an action (e.g., mime the word 'hit'), not just when they speak a word out loud. Based on several lines of research I expect to find a production effect for mimed words that is as strong, or stronger in comparison with spoken words.

The reason I didn't tell you these hypotheses at the start of the study is because I didn't want to influence your memory strategies. To be able to accurately interpret the results I need to be sure that you were trying to remember all the words presented, regardless of whether you spoke, mimed or read them silently.

Now that you know what the hypotheses of the study are, I would like to ask whether you still consent your data being included in the study. As discussed at the start of the study you are free to withdraw your data if you choose. Do you have any questions?

[Experimenter will answer any questions. If the participant requests to withdraw their data the experimenter will delete the data from the computer and destroy the signed consent form by tearing it up.

Once all the participant's questions have been answered and inclusion of data confirmed (or withdrawn), the experimenter will thank the participant for their participation and end the study session.]